

## CHAPITRE 5

CÉLINE DUGUA, MARIE BACLESSE

### Incidence d'effets de fréquence sur l'usage de la liaison en lecture à haute voix et dans des jugements normatifs chez des enfants de CE2-CM1

#### 1. Introduction

##### 1.1 *La liaison dans les différentes modalités langagières*

L'usage et le traitement de la liaison est aujourd'hui bien documenté et ce dans les différentes modalités langagières. En production chez l'adulte, on dispose de corpus oraux qui permettent de circonscrire les usages en fonction de différentes situations de communication. Rappelons (1) le corpus PFC<sup>1</sup>, dont les protocoles ont été élaborés dans le but de décrire cette particularité de la langue française ; (2) les travaux plus anciens de De Jong (Booij & De Jong, 1987 ; De Jong, 1994) sur le corpus d'Orléans, qui rendent compte, dans le cadre d'entretiens semi-directifs, de l'usage de la liaison en fonction de facteurs sociaux tels que le sexe, l'âge et la catégorie socio-professionnelle des locuteurs ; (3) le corpus d'Encrevé (1988) sur les discours de 21 hommes politiques dans les années 1978-1981. Évidemment, nous ne reviendrons pas sur l'intérêt que représente le travail à partir de données « naturelles » pour observer le phénomène de la liaison, connu pour être socialement marqué. Soulignons que les contraintes méthodologiques, telles que le temps de transcription et de codage des liaisons, contribuent à rendre ce type de don-

---

1 <[www.projet-pfc.net](http://www.projet-pfc.net)> et chapitres 2 et 3, ce volume

nées relativement marginal. Malgré ces contraintes, de plus en plus de travaux ont la volonté de rendre compte de l'usage réel de la liaison chez de jeunes locuteurs à partir de situations naturelles, citons Basset (2000) qui a observé la liaison chez des enfants de 3, 7 et 11 ans dans leurs activités quotidiennes (repas, jeux, bain), Méradji et Grégoire (2001) ont travaillé avec des jumeaux de 5;8 lors d'interactions familiales, Nardy (2008) a recueilli des données dans des salles de classe de maternelle, Chabanal et Liégeois (ce volume) ont mis en place une méthodologie de recueil de corpus denses avec 4 enfants entre 28 et 48 mois. L'ensemble de ces données fournit aujourd'hui un panorama riche du développement de la liaison chez les jeunes enfants et de son usage entre 2 et 10 ans. Il en ressort (pour plus de détails, voir notamment Chevrot *et al.* (2007)) que, jusqu'à l'âge de 4 ans, les enfants produisent fréquemment des formes erronées (ex : *les ours* produit /letuʁs/ avec un /t/ à la place du /z/, *un petit escargot* produit /œptineskaʁgo/ avec un /n/ à la place du /t/, *un arbre* produit /œaʁbʁ/ sans consonne de liaison), puis à partir de 4 ans, les liaisons dites obligatoires sont quasiment acquises. Quant aux liaisons facultatives (tous contextes confondus), elles sont rarement réalisées jusqu'à environ 6 ans, puis, à partir de cet âge, atteignent un score comparable aux scores obtenus chez des adultes en situation non formelle (Dugua & Spinelli, 2009).

Du point de vue du traitement cognitif en perception, l'une des questions à laquelle les chercheurs s'intéressent est de savoir si la présence d'une liaison constitue un coût de traitement perceptif. Depuis Yersin-Besson et Grosjean (1996) jusqu'aux travaux menés par Spinelli et collègues (Spinelli, Cutler & Mcqueen, 2002 ; Spinelli, Mcqueen & Cutler, 2003 ; Spinelli & Meunier, 2005), les résultats expérimentaux tendent à montrer que la présence d'une liaison n'entraîne pas un coût de traitement supplémentaire par rapport à un contexte d'enchaînement<sup>2</sup> ou de

---

2 L'enchaînement consiste à rapprocher deux mots graphiques en reliant le dernier phonème consonantique du mot1 (phonème toujours prononcé) au premier phonème vocalique du mot2. Par exemple, il y a enchaînement dans « il arrive » /ilaviv/, alors qu'il y a liaison dans « vous arrivez » /vuzavive/. Dans le cas de l'enchaînement, la consonne /l/ de « il » est toujours prononcée, même quand le mot est dit en isolation, alors que dans le cas de la liaison, le /z/ n'est pas prononcé

correspondance entre frontières syllabique et lexicale. Gaskell *et al.* (2002) trouvent d'ailleurs que la liaison et l'enchaînement faciliteraient la reconnaissance du mot2. Il semblerait que l'apparition de la consonne de liaison agisse comme le signe qu'un mot à initiale vocalique va suivre. La perception s'appuierait donc davantage sur la reconnaissance du mot précédant la consonne resyllabée que sur la resyllabation elle-même. Ainsi, semble-t-il, le système de reconnaissance utiliserait l'information lexicale spécifiant qu'un mot1 est déclencheur de liaison pour identifier la consonne finale comme étant resyllabée au mot2. Des expériences menées par Spinelli (Spinelli *et al.*, 2002) suggèrent le recours à des facteurs lexicaux pour résoudre la resyllabation inhérente à la liaison. Ces auteurs montrent par exemple que lorsque les sujets sont confrontés à des contextes de liaison illégaux (ex : *demi-t-agneau*), dans une tâche de décision lexicale, le temps de reconnaissance est plus long que dans de réels contextes de liaison (ex : *petit agneau*). Ainsi la nature lexicale du mot1 (mot1 pouvant faire liaison ou pas) serait utilisée dans la reconnaissance du mot2. Enfin, outre les facteurs lexicaux, il semblerait que la longueur de la consonne de liaison (consonne plus brève qu'une consonne initiale de mot) constitue une caractéristique utile pour distinguer les mots2 à initiale vocalique et les mots2 à initiale consonantique.

Production orale et perception sont, historiquement, les deux premiers domaines d'études en ce qui concerne la question de la liaison. Plus récemment, notamment grâce à l'ANR Phonlex, les chercheurs se sont intéressés à la question du traitement de la liaison à l'écrit et des difficultés que ce phénomène oral peut causer pour des enfants en phase d'apprentissage. Nous citerons évidemment les travaux de Soum Favaro et Gunnarsson (Gunnarsson & Soum-Favaro, 2008 ; Soum Favaro, Gunnarsson & Largy, ce volume) qui montrent que chez les enfants de 7-10 ans (CE2-CM2), l'écriture des contextes de liaison (*un gros avion*) entraîne des erreurs – les auteurs relèvent d'ailleurs davantage d'erreurs dans les contextes de liaison qu'en contextes de mots2 à initiale conso-

---

quand « vous » est produit en isolation. C'est le contexte de liaison qui permettra la manifestation du phonème /z/.

nantique (ex : *un gros zèbre*) – et que les enfants sont sensibles à des paramètres phonético-acoustiques.

Enfin, en lecture, il n'existe pas d'études portant spécifiquement sur les réalisations des liaisons. Le corpus PFC comprend pourtant une tâche de lecture de texte et une tâche de lecture de mots, mais, à notre connaissance, ces deux tâches n'ont pas fait l'objet de recherches spécifiques. Durand et Lyche (2008) expliquent ci-dessous les raisons de la présence de ces deux tâches de lecture dans le protocole PFC :

The reading tasks were integrated within the protocol for two reasons: (i) they guarantee a full comparability of the results and (ii) they are required once the phonological goals have been clearly formulated. No oral corpus, regardless of its size, can claim to include all the data the linguist is looking for, but the reading tasks give us systematic access to much of the phonological information we seek and, in addition, to a formal register. They also allow the testing of hypotheses concerning the relationship between speech and writing. (Durand & Lyche, 2008 : 37-38)

En effet, la lecture à haute voix permet un accès aux informations phonologiques. Il convient toutefois de ne pas négliger l'influence que peut avoir l'accès aux formes graphiques sur certaines réalisations phonologiques, notamment sur la liaison. Signalons, toutefois, que chez des enfants de primaire, Soum-Favaro (2010) note que les résultats sont meilleurs dans une tâche de dictée qu'en lecture à haute voix. L'accès à la forme graphique des mots ne serait donc pas un gage de réussite en production de liaison.

## 1.2 *Enjeu de notre recherche*

Travailler sur la lecture à haute voix en tant que modalité orale qui s'appuie sur de l'écrit constitue le principal enjeu de ce travail. On sait que phonie et graphie s'influencent mutuellement. Dès les années 1930, Buben (1935) a formalisé ces principes et a donné son nom au phénomène d'influence de la graphie sur la phonie (*effet Buben*) et de la phonie sur la graphie (*contre effet Buben*). Plus récemment, le modèle interactif bimodal de la reconnaissance des mots parlés et écrits (Ferrand, 2001 ; Grainger & Van Heuven, 2003 ; Ziegler, Ferrand & Montant, 2004 ; Ziegler, Muneaux & Grainger, 2003), illustre la très forte interdépen-

dance de ces deux modules. La donnée phonologique informe la reconnaissance des mots écrits, et l'information orthographique influence la reconnaissance des mots parlés. Cependant, ce modèle postule que, dans la mesure où le langage parlé prime sur le langage écrit (au niveau du développement et de l'espèce humaine), les connexions bidirectionnelles entre la phonologie et l'orthographe sont asymétriques : l'influence du langage parlé sur le langage écrit est supposée plus forte que celle du langage écrit sur le langage parlé. La liaison, qui se manifeste à l'oral par la production d'un son consonantique, constitue ainsi un phénomène très intéressant pour illustrer les influences réciproques entre oral et écrit.

Notre travail se situe à la périphérie de ces questions puisque nous nous interrogeons sur le rapport entre langage parlé et lecture, en prenant appui sur les principes des théories basées sur l'usage. Dans leur introduction de l'ouvrage *Usage-based models of language use*, Kemmer et Barlow (2000) présentent, sous la forme de neuf principes, les conceptions partagées par l'ensemble des recherches s'inscrivant dans le cadre de ces théories. Celles-ci ont trouvé un large écho dans les approches développementales de l'acquisition du langage (Bates & Goodman, 1997 ; Tomasello, 2003), l'une des raisons étant que la notion d'usage renvoie aux différentes sources d'input langagier et aux interactions à partir desquelles l'enfant construit son système linguistique. Selon cette conception, le système linguistique du locuteur est fondé sur les « événements d'usage », c'est-à-dire, les énoncés concrets que le locuteur produit et entend. En ce sens, les représentations linguistiques sont étroitement liées aux événements d'usage puisque ces derniers constituent la base à partir de laquelle le système linguistique du locuteur se forme et évolue. Directement associée à l'usage, la notion de fréquence des occurrences dans l'input va jouer un rôle essentiel à la fois en ce qui concerne les processus d'acquisition par les enfants mais aussi, plus généralement, dans le traitement cognitif adulte. Dans les théories basées sur l'usage, le système étant piloté par les expériences du locuteur et leur accumulation, la fréquence des instances apparaît comme un facteur premier de sa structure et de son fonctionnement (Kemmer & Barlow, 2000).

Or, nous avons pu montrer qu'à l'oral la fréquence de réalisation de la liaison chez l'adulte influence l'usage de la liaison chez les jeunes enfants (Dugua, Spinelli, Chevrot & Fayol, 2009). A partir d'une tâche

de dénomination d'images, où la moitié des mots est le plus souvent employée au singulier (ex : *ours*) et l'autre moitié plus souvent au pluriel (ex : *arbre*), proposée à 122 enfants âgés entre 3 et 6 ans dans les contextes *un* + X et *deux* + X, nous observons que plus un nom est fréquemment employé au pluriel (*arbre*), plus les enfants de 3-4 ans (les plus jeunes enfants de l'échantillon) réalisent correctement la liaison en /z/ après *deux*, et plus la liaison en /z/ intervient dans les erreurs par remplacement après *un* (erreurs de type : /œzaʁbœ/). Au-delà de 4 ans, on ne trouve plus ces corrélations.

Dans ce chapitre, nous nous demanderons si les effets de fréquence de réalisation de la liaison à l'oral affectent également leur réalisation chez des enfants de CE2-CM1 dans trois types de tâches, deux en modalité orale (dénominations d'images et jugements normatifs) et une au croisement de la modalité orale et de la modalité écrite : la lecture à haute voix.

Le contrôle de la fréquence d'usage fera l'objet de deux types d'observations. D'une part, il permettra de distinguer les productions de liaisons obligatoires (désormais LO) et de liaisons facultatives (désormais LF) : les LO étant les liaisons qui sont systématiquement réalisées, quelles que soient la situation de communication et les caractéristiques du locuteur (Booij & De Jong, 1987) ; les LF sont alors celles qui présentent de la variation. D'autre part, et de façon plus fine, nous contrôlerons la fréquence de réalisation des liaisons au sein des LF en diversifiant les contextes grammaticaux, depuis ceux dans lesquels les liaisons sont très rarement réalisées (après *depuis*, après *pendant*, etc.) jusqu'à ceux dans lesquels les liaisons sont très souvent réalisées (après *en*, après *très*, etc.). Ainsi, à partir du travail de De Jong (1994), nous établirons un continuum des usages adultes (Figure 1) qui nous servira de base pour analyser les productions enfantines dans les tâches de jugements normatifs et de lecture à haute voix.

## 2. Méthodologie

### 2.1 Participants

Cent onze enfants âgés entre 8 et 11 ans (moyenne : 9;2), qui se répartissent entre les classes de CE2 (66 enfants : 38 filles et 28 garçons) et de CM1 (45 enfants : 25 filles et 20 garçons), ont participé à l'étude. Le niveau de lecture de ces enfants n'a pas été contrôlé, nous nous assurons seulement qu'ils n'avaient pas et n'avaient pas eu de suivi orthophonique. Nous avons choisi des enfants de cycle 3 car, d'après le bulletin officiel<sup>3</sup>, l'élève doit parvenir à « une automatisation de la reconnaissance des mots, ce qui doit rendre aisée la lecture de mots irréguliers et rares et augmenter la rapidité et l'efficacité de la lecture (silencieuse) ». Sprenger-Charolles *et al.* (2003) ont par ailleurs montré dans une étude longitudinale qu'entre la fin du CE1 et la fin du CM1, la lecture de mots réguliers et de pseudomots entraîne moins de 20% d'échec et que c'est à partir de la fin du CE1 que la proportion d'échec se stabilise. Choisir des enfants de CE2-CM1 permet de s'assurer qu'ils maîtrisent suffisamment la lecture de mots pour être capables de les enchaîner et éventuellement de réaliser des liaisons, étant bien conscient, cependant, que la variabilité interindividuelle en ce qui concerne le « niveau de lecture » des enfants de CE2-CM1 est forte.

### 2.2 Protocole

Trois tâches expérimentales ont été proposées aux enfants – dénomination d'images, jugement normatif, lecture à haute voix – l'ensemble des passations a été enregistré. Nous avons utilisé les mêmes énoncés dans les tâches de lecture et de jugement, énoncés qui ont été présentés aléatoirement à l'enfant. La tâche de dénomination, quant à elle, s'appuie sur un matériel linguistique différent. Elle permettra simplement d'observer une éventuelle différence entre l'usage des LO et des LF, tout comme on

---

3 <[www.education.gouv.fr/bo/2008/hs3/programme\\_CE2\\_CM1\\_CM2.htm](http://www.education.gouv.fr/bo/2008/hs3/programme_CE2_CM1_CM2.htm)>

le constatait chez des enfants plus jeunes. Dans cette tâche de dénomination d'images, tous les enfants ont commencé par produire les LO, puis les LF.

### 2.2.1 Tâche de dénomination d'images

La tâche de dénomination d'images s'inspire de celles généralement utilisées avec des enfants plus jeunes (Dugua, 2006 ; Nardy, 2008). Elle comprend deux types de liaison : les liaisons obligatoires dans les contextes *un* + X et *deux* + X et les liaisons facultatives dans les contextes *petit* + X et *gros* + X. Pour faire produire les premières, nous représentons les images en un et deux exemplaires, pour les secondes en petit et gros exemplaires. Pour les deux types de liaison, six mots-cibles ont été utilisés : *arbre*, *avion*, *écureuil*, *éléphant*, *ordinateur* et *ours*, et six mots distracteurs : *camion*, *canard*, *chat*, *citron*, *téléphone* et *tigre*. Chaque présentation d'une image représentant un mot cible alternait avec celle représentant un mot distracteur.

Cette tâche permettra de nous assurer que les LO entre déterminant et nom ne posent plus de difficultés aux enfants de CE2-CM1. On s'attend en revanche à relever de la variation dans la production des LF, liaisons qui se caractérisent par de la variation dans l'usage adulte.

### 2.2.2 Tâche de lecture à haute voix

La tâche de lecture à haute voix a été construite autour de 44 énoncés, indépendants les uns des autres, imprimés sur des cartes que l'enfant tirait au sort l'une après l'autre. Le caractère aléatoire de l'ordre de présentation était donc assuré.

Parmi ces 44 énoncés, 12 comprenaient des liaisons obligatoires en contextes *un* + X et *deux* + X (les mêmes que dans la tâche de dénomination). Les 32 énoncés en contexte de liaison facultative ont été créés à partir des travaux de De Jong (1994). Cet auteur a constitué un sous-corpus du corpus d'Orléans<sup>4</sup> (Blanc & Biggs, 1971), sous-corpus comprenant 45 entrevues informelles d'environ une heure, enregistrées entre 1969 et 1970 avec des locuteurs adultes. Les témoins étaient répartis en

---

4 Ce sous-corpus comprend 16 000 contextes de liaisons possibles.



cinq catégories socioéconomiques et trois groupes d'âge (18-29, 30-49, 50 et plus). De Jong (1994) a sélectionné trois grandes catégories linguistiques : les verbes, les prépositions et les adverbes. Au sein de ces catégories, il a affiné ses observations en relevant les pourcentages de liaisons réalisées selon les mots<sup>1</sup> impliqués. Par exemple, dans la catégorie « après préposition », il détaille « après *chez* », « après *sans* », etc. Il donne non seulement le taux de réalisation de la liaison dans chacun des contextes lexicaux, mais aussi le nombre de contextes de liaisons relevées dans son sous-corpus.

Au sein de chacune des trois grandes catégories, nous avons sélectionné les cinq contextes qui présentent le nombre d'occurrences le plus important afin de nous appuyer sur des résultats tangibles, à savoir : (1) après l'auxiliaire *avoir*, (2) après des auxiliaires modaux, (3) après des adverbes, (4) après l'auxiliaire *être*, (5) après des prépositions<sup>5</sup>. Pour les trois premiers contextes, nous avons choisi deux formes lexicales dont la fréquence de réalisation était contrastée, et pour les contextes (4) et (5), quatre formes lexicales ont été retenues : *est*, *sont*, *suis*, *était* pour l'auxiliaire *être* ; pour les prépositions deux monosyllabiques *en* et *chez* et deux polysyllabiques *depuis* et *pendant*. Pour ces derniers, nous avons également veillé à contraster les fréquences de réalisation. Par exemple, dans la catégorie des « auxiliaires modaux », nous avons retenu les formes qui présentaient à la fois le plus grand nombre d'occurrences (94 pour *peut* et 88 pour *faut*) et le pourcentage de réalisation de liaison le plus contrasté (respectivement 20.2% et 4.5%). Ainsi, pour chaque catégorie, nous avons conservé deux formes : une dont la fréquence de réalisation de la liaison était élevée, et une autre dont le pourcentage était moins élevé. Par ailleurs, dans le choix des formes lexicales, nous avons veillé à ce que les différents taux de réalisation se répartissent régulièrement sur un continuum entre des liaisons jamais réalisées (après *depuis*) et des liaisons qui s'apparentent à des LO (après *très*), comme l'illustre la figure ci-après.

---

5 Ces terminologies sont celles utilisées par l'auteur.

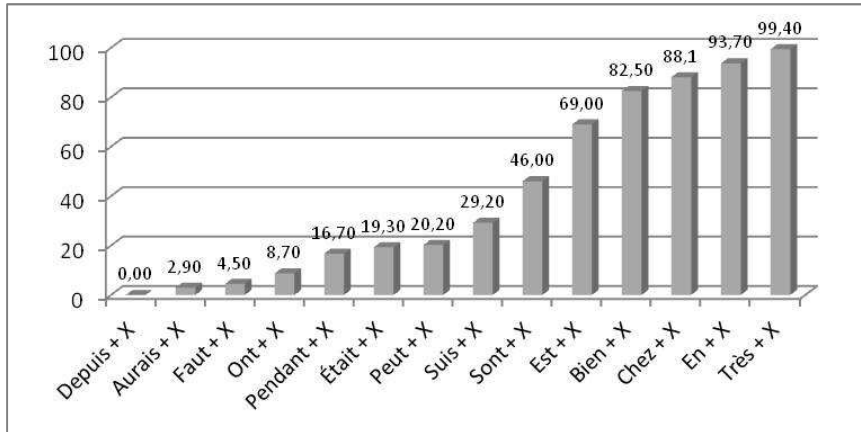


Figure 1 - Fréquence de réalisation des liaisons en fonction du contexte dans les données de De Jong (1994)

Dans un deuxième temps, nous avons déterminé le terme qui constituerait le mot2 du contexte de liaison. Afin de contrôler un éventuel effet des co-occurrences mot1-mot2 sur l'usage de la liaison, nous avons recherché les fréquences de ces dernières dans la base de données Lexique 3<sup>6</sup> (New, Pallier, Ferrand & Matos, 2001), et plus particulièrement dans le corpus de sous-titres de films, étant le corpus le plus proche d'un usage oral. Pour chaque contexte de liaison, nous avons constitué deux énoncés : l'un avec une fréquence de co-occurrence mot1-mot2 forte, et l'autre avec une fréquence de co-occurrence faible, et pour quatre contextes (*est + X*, *était + X*, *suis + X*, *très + X*), nous avons créé un troisième énoncé dont la fréquence de co-occurrence se situait entre les co-occurrences forte et faible.

En bref, pour créer les énoncés, la fréquence a été contrôlée à deux niveaux :

Premier niveau : le contrôle des fréquences de réalisation des liaisons en fonction du contexte et de la nature lexicale du mot1 (cf les travaux de De Jong (1994))

---

6 <www.lexique.org>

Second niveau : le contrôle des fréquences de co-occurrences mot1-mot2 : une co-occurrence fréquente et une co-occurrence rare pour chaque forme de mot1 choisie (cf. la base de données Lexique 3)

Nous pouvons alors représenter ce double contrôle par le schéma suivant :

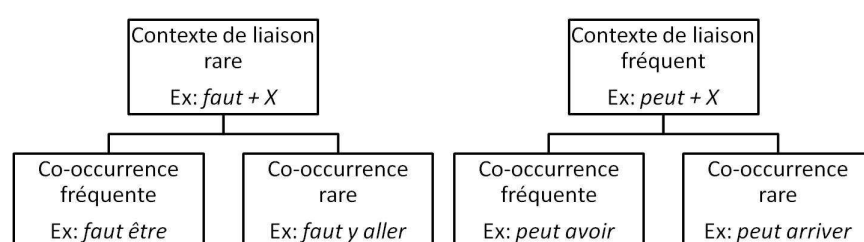


Figure 2 - Représentation du double niveau de contrôle pour la constitution des énoncés à lire et à juger (exemple avec la catégorie « auxiliaires modaux »)

### 2.2.3 Tâche de jugement normatif

Deux objectifs ont guidé l'élaboration de la tâche de jugement normatif. D'une part, il s'agissait de mesurer la connaissance de la norme qu'ont les enfants en ce qui concerne l'usage de la liaison facultative. D'autre part, nous cherchions à déterminer si cette connaissance de la norme pouvait être guidée par la fréquence d'usage des contextes de liaison.

Dans cette tâche, l'enfant doit juger des paires de séquences en contexte de liaison obligatoire et facultative, ces paires comprenant un énoncé réalisé avec liaison et l'autre réalisé sans liaison. Il s'agit des mêmes items que pour l'épreuve de lecture à haute voix, qui répondent donc aux mêmes critères de sélection. Cependant, il ne nous a pas été possible d'utiliser, pour chaque enfant, toutes les séquences de la tâche de production<sup>7</sup>. Nous avons donc fait le choix de créer deux jeux de cartes pour réduire le nombre de séquences à juger. La moitié des enfants a jugé le jeu de cartes 1 et l'autre moitié le jeu de cartes 2. Ainsi, pour chaque jeu

<sup>7</sup> Si nous avions intégré à l'épreuve de jugement tous les items, nous aurions fait juger à l'enfant 44 énoncés. Cela aurait représenté une expérience trop coûteuse en temps et en attention.

de cartes, nous avons conservé tous les énoncés impliquant une liaison facultative (soit 32), car l'effet de fréquence paraissait plus pertinent à observer sur celles-ci. En revanche, pour les liaisons obligatoires, nous avons organisé les mots<sub>2</sub> en deux groupes que nous avons contrebalancés avec les mots<sub>1</sub>. Le groupe 1 était constitué des mots<sub>2</sub> *arbre*, *escargot*, *ordinateur* et le groupe 2 des mots<sub>2</sub> *ours*, *avion*, *éléphant*. Dans le premier jeu de cartes, les mots du premier groupe étaient jugés au singulier après le mot<sub>1</sub> *un* et les mots du deuxième groupe jugés au pluriel après le mot<sub>1</sub> *deux* et inversement pour l'autre jeu de cartes.

Par ailleurs, outre cette structuration pour les LO, les deux jeux de cartes se distinguent par l'ordre de présentation des deux formes de chaque énoncé (en LO et en LF). En effet, afin d'éviter un biais méthodologique lié à l'ordre de présentation des énoncés, nous alternons la présentation de ces formes entre les deux jeux de cartes. Par exemple, si dans le premier jeu de cartes, l'énoncé « *Il s'est bien adapté* » est présenté d'abord sous la forme réalisée « *Il s'est bien Nadapté* », puis sous la forme non réalisée « *Il s'est bien Oadapté* », dans le deuxième jeu de cartes, ce même énoncé est présenté dans l'ordre inverse, sous la forme non réalisée d'abord « *Il s'est bien Oadapté* », puis sous la forme réalisée « *Il s'est bien Nadapté* »<sup>8</sup>.

Les cartes sont présentées aléatoirement pour chaque enfant. On demande à l'enfant de déterminer, parmi les deux formes proposées sur chaque carte, laquelle est « la mieux dite ». Cette consigne induit donc une référence à la norme prescriptive.

### 3. Résultats

Pour plus de clarté, nous présentons les résultats par modalités langagières.

---

<sup>8</sup> Le codage utilisé ici indique à l'enquêteur comment produire les énoncés. Dans cette tâche, les enfants n'ont pas accès aux cartes mais seulement à la lecture qu'en fait l'enquêteur. Une consonne de liaison en majuscule devant le mot<sub>2</sub> signifie que la liaison doit être réalisée, le signe « *O* » signifie qu'elle ne doit pas l'être.

### 3.1 Tâche de dénomination d'images

Des précédentes études ont montré une influence de la fréquence sur les productions et la reconnaissance des mots en contexte de liaison obligatoire (Dugua, 2008 ; Dugua *et al.*, 2009). Dans ces études, la fréquence était mise en évidence par l'orientation en nombre des noms. Nous avons pu repérer que certains noms sont majoritairement employés au singulier, et d'autres au pluriel. Ces caractéristiques semblent déterminer les types de productions des enfants de 3-4 ans, à la fois en termes de productions justes et d'erreurs par remplacement, ainsi que la reconnaissance des mots chez les enfants de 5-6 ans.

Les résultats de la présente tâche de dénomination vont permettre de s'assurer que les enfants de CE2-CM1 maîtrisent les LO entre déterminant et nom. Nous nous attendons à ce résultat puisque les LO sont celles qui ne présentent pas de variation dans la parole adulte. En d'autres termes, leur fréquence de réalisation se situe à 100%, contrairement aux LF où on attend de la variation dans les productions.

#### 3.1.1 Dénomination des liaisons obligatoires entre déterminant et nom

Nous relevons que les liaisons obligatoires sont produites correctement dans 98% des cas (ex : l'enfant dit *deux ours* /døzuxs/ avec la consonne /z/ attendue), les autres productions sont donc très marginales : 1.85% d'omission (cas où la liaison n'est pas réalisée, l'enfant dit *deux ours* /døuxs/) et 0.15% de remplacement (cas où la consonne produite n'est pas celle attendue, l'enfant dit par exemple /dønuxs/ ou /døtuxs/).

Comme nous l'attendions, on peut donc considérer que les enfants de CE2-CM1 maîtrisent la liaison obligatoire entre déterminant et nom. Comparativement aux résultats obtenus chez des enfants de 5-6 ans sur une tâche similaire (Dugua, 2006), on constate que ce sont les erreurs par omission de liaison qui ont le plus évolué, passant de 13.9% des productions à 1.85%.

### 3.1.2 Dénomination des liaisons facultatives entre adjectif et nom

Rappelons que le contexte choisi pour observer la production orale de liaisons facultatives (entre adjectif et nom) est un contexte à fort taux de réalisation, puisque De Jong (1994) relève 94.3% de réalisation dans les entretiens avec des adultes. Nous nous attendons donc à des réalisations importantes chez les enfants, cependant moindre que pour les LO.

Dans notre échantillon d'enfants, nous relevons 50.8% de liaisons réalisées, 34.6% de liaisons non réalisées et 14.6% d'erreurs par remplacement. La forme avec liaison est significativement plus utilisée que la forme sans liaison ( $t(110) = 3.531, p < .001$ ), même si elle reste bien inférieure à la valeur de De Jong chez des adultes en situation d'entretien.

Ces résultats généraux cachent des différences fortes entre les deux mots1 utilisés (*petit* et *gros*), comme nous pouvons le constater dans le Tableau 1.

Tableau 1- Production des LF en fonction du mot1

	Après <i>petit</i>	Après <i>gros</i>
Taux de liaisons réalisées	84.9%	15.9%
Taux de liaisons non réalisées	14.8%	54.8%
Taux d'erreurs par remplacement	0.3%	29.3%

La réalisation des liaisons est quasi systématique après *petit* (84.9%) et il n'y a presque pas d'erreurs par remplacement. En revanche, la répartition est bien différente après *gros*. La forme la plus fréquente est la non réalisation de la liaison (54.8%), puis les erreurs par remplacement (29.3%) et enfin les liaisons réalisées (15.9%). On a donc deux comportements bien distincts entre ces deux mots que l'on peut expliquer à différents niveaux. L'essentiel des erreurs par remplacement après *gros* (97%) se manifeste par la substitution du /z/ en /t/ (ex : *gros arbre* prononcé /gʁotɑʁbɛ/). On peut penser que ces erreurs résultent d'un calque opéré avec l'adjectif *grand* qui induit une liaison en /t/ et qui est beaucoup plus fréquent que *gros*. Par ailleurs, un autre effet de fréquence peut expliquer

ces différences : l'adjectif *gros* est plus rare que l'adjectif *petit*<sup>9</sup>, la probabilité de le rencontrer en contexte de liaison est donc moindre. Cette caractéristique peut expliquer que son usage est moins stabilisé que celui de l'adjectif *petit* chez les enfants de notre échantillon.

### 3.2 Tâche de jugement normatif

La tâche de jugement normatif a été mise en place afin de vérifier le degré de connaissance de la norme prescriptive en ce qui concerne l'usage des liaisons facultatives qu'ont les enfants de CE2-CM1. D'autre part, au sein des différents contextes de liaisons facultatives utilisés, différents de par leur fréquence de réalisation chez les locuteurs adultes, nous nous attendons à trouver une corrélation entre les jugements vers la forme liaisonnée et le taux de réalisation des liaisons des contextes considérés. En effet, comme l'a souligné Nardy (2008: 183) :

En psycholinguistique, et plus particulièrement selon la théorie des exemplaires et les théories basées sur l'usage, la connaissance linguistique se construit à partir des événements d'usage perçus et produits. Il est donc attendu que les jugements soient sensibles à des effets de familiarité avec certaines séquences. En effet, les séquences les plus fréquentes dans l'input et l'output devraient être plus disponibles et, en conséquence, être reconnues comme familières et être évaluées plus favorablement que des séquences moins fréquentes.

Les jugements en faveur de la forme de LO réalisée sont de 94.6%, et 5.4% en faveur des LO non réalisées, cette différence est significative ( $t(111) = 37.7, p < .001$ ).

Pour les contextes de LF également, les jugements en faveur de la forme liaisonnée des énoncés sont significativement plus importants (60.9%) que les jugements en faveur de la forme non liaisonnée (39.1%) ( $t(110) = 5.3, p < .001$ ). Alors qu'on aurait pu attendre sur ce type de tâche des choix plus marqués vers la forme liaisonnée, notamment en raison du protocole lui-même (toujours opposer une forme liaisonnée à

---

9 D'après la base de données Lexique 3 (consulté le 20/02/2012), dans les corpus « Livres » et « Films », l'adjectif *gros* apparaît 396.92 fois par million, l'adjectif *grand* 876.24 et l'adjectif *petit* 1227.5

une forme non liaisonnée), subsiste une variation importante puisque plus du tiers des réponses ont été des choix vers la forme non liaisonnée.

Cette tâche comprenait 32 énoncés à juger en LF et 6 énoncés en LO, les contextes de liaison ayant été choisis en fonction de la fréquence de réalisation de ces liaisons chez des locuteurs adultes (De Jong, 1994). Pour chaque contexte, nous disposons donc du taux de réalisation chez les adultes, et du taux de jugements vers la forme liaisonnée chez notre échantillon d'enfants. Le graphique de la figure 3 permet de visualiser, pour chaque contexte utilisé, la fréquence de réalisation chez De Jong (en gris clair) et les taux de jugement en faveur de la forme liaisonnée chez les enfants (en gris foncé).

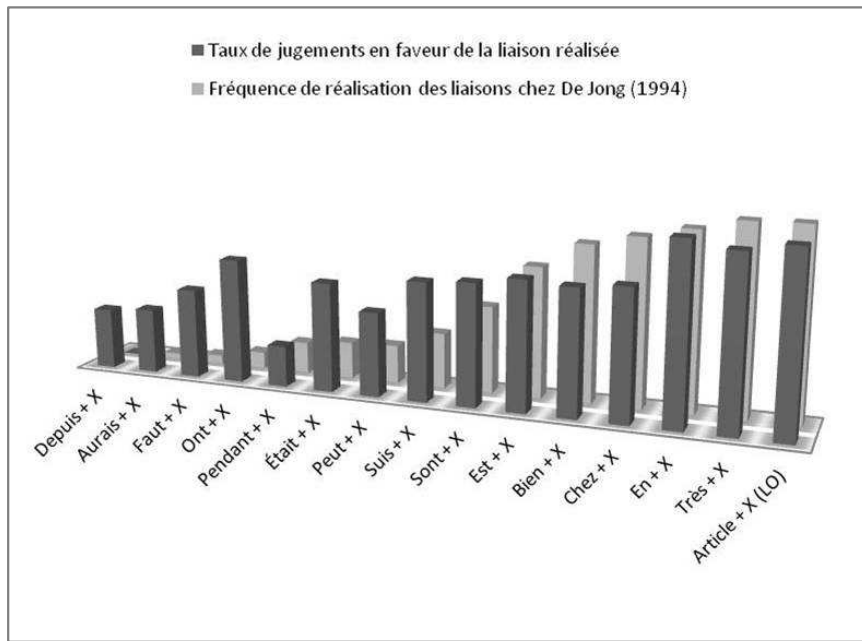


Figure 3 – Résultats de la tâche de jugement en LO et LF et données de De Jong (1994) en fonction du contexte



Nous avons appliqué un test de corrélation<sup>10</sup> (corrélation de Pearson), dont les résultats font apparaître une corrélation positive et significative entre les deux données ( $r = .813$ ,  $p < .001$ ). Ce résultat signifie que les enfants ont davantage tendance à choisir la forme liaisonnée des séquences dans lesquelles la liaison est fréquemment réalisée chez l'adulte (ex : choisir *très Zembêtant* plus souvent que *très Oembêtant*) et inversement (choisir *depuis Oun moment* plus souvent que *depuis Zun moment*).

### 3.3 Tâche de lecture à haute voix

En ce qui concerne la tâche de lecture à haute voix, nous formulons deux hypothèses toutes deux régies par la question de la fréquence. Premièrement, nous nous attendons à ce que les LO (les liaisons systématiquement réalisées par les adultes) soient fortement réalisées, comme elles le sont en production orale. On ne s'attend pas à ce que la lecture de liaisons pose des problèmes aux enfants sur les LO. Deuxièmement, nous nous attendons à une moindre réalisation des LF (celles dont la réalisation varie chez l'adulte) et plus précisément que les taux de réalisation des LF varient en fonction de la réalisation de ces LF dans la parole adulte (De Jong, 1994) et en fonction de la fréquence de co-occurrence mot1-mot2 obtenue grâce à la base de données Lexique 3.

En lecture d'énoncés, il apparaît que 96.5% des liaisons obligatoires sont correctement réalisées. Subsistent 3.1% d'erreurs par omission et 0.4% d'erreurs par remplacement.

Globalement, chez les enfants de notre échantillon, les liaisons facultatives réalisées représentent 31.9% des productions, et les non réalisées 68.1%. Cette différence est significative ( $t(110) = -9.775$ ,  $p < .0001$ ).

Les résultats globaux ont été complétés par des résultats en fonction des contextes de liaison. Le graphique ci-après permet de visualiser les taux de réalisations des liaisons en fonction du contexte, organisé selon l'ordre croissant des valeurs chez De Jong.

---

10 La corrélation a été appliquée uniquement sur les résultats de LF.

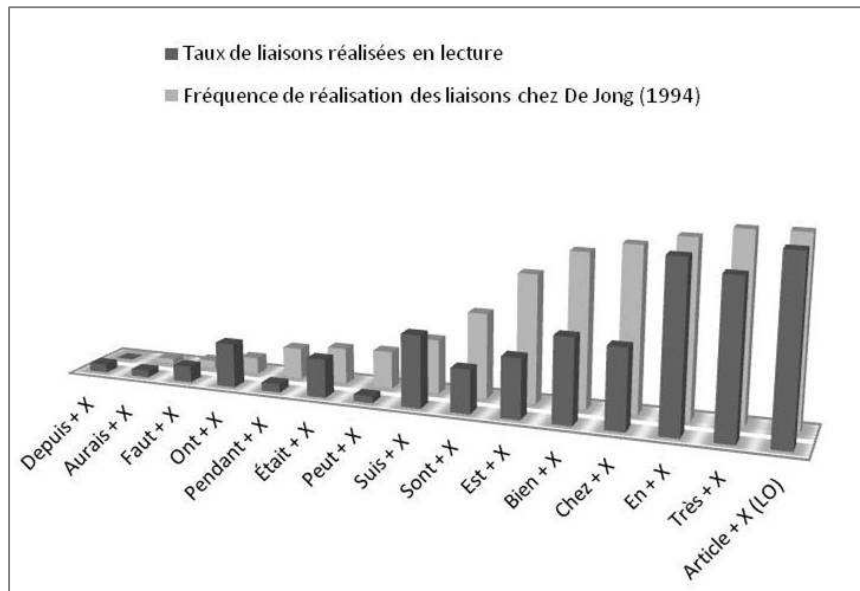


Figure 4 - Résultats de la tâche de lecture en LO et LF et données de De Jong (1994) en fonction du contexte

En comparant les moyennes des taux de réalisation chez les enfants et les données de De Jong en fonction des 14 contextes de liaisons facultatives sélectionnés, nous relevons que les adultes réalisent significativement plus la liaison en situation d'entretien oral (41.4%) que les enfants de CE2-CM1 en situation de lecture (30.4%) ( $t(13) = 2.197$ ,  $p = .047$ ). D'autre part, le nombre de contextes pour lesquels les enfants réalisent davantage la liaison que les adultes est limité (6 contextes sur 14 dont 2 se distinguent par moins de 1%). Ces contextes se situent tous dans la partie gauche du graphique, c'est-à-dire des contextes pour lesquels les adultes réalisent peu la liaison (moins de 30% de réalisation).

Par ailleurs, les résultats du test de la corrélation de Pearson font apparaître une corrélation positive et significative entre la réalisation des LF et les taux de liaison chez les adultes ( $r = .872$ ,  $p < .0001$ ). En revanche, on ne trouve pas de corrélation entre le taux de réalisation des

LF en lecture et la fréquence de co-occurrence mot1-mot2 dans Lexique 3 ( $r = -.051$ ,  $p = .782$ ).

### 3.4 Tâches de lecture et de jugement

Rappelons les résultats globaux obtenus dans les deux modalités de jugement et de lecture en ce qui concerne les liaisons facultatives.

Tableau 2 - Résultats des LF en tâches de lecture et de jugement

	Tâche de jugement	Tâche de lecture
Taux de liaisons réalisées	60.9%	31.9%
Taux de liaisons non réalisées	39.1%	68.1%

Le taux de liaisons jugées réalisées est significativement plus important que le taux de liaisons lues réalisées ( $t(110) = -12.930$ ,  $p < .0001$ ). Globalement, les enfants jugent mieux qu'ils ne produisent.

## 4. Conclusion et discussion

Sans revenir sur l'ensemble des résultats obtenus dans cette étude, nous retiendrons que les enfants de CE2-CM1 maîtrisent l'usage de la liaison obligatoire entre déterminant et nom en dénomination (98% de LO correctes). En tâche de lecture à haute voix, la réalisation des LO est forte (96.5%) et les jugements normatifs se tournent le plus souvent vers la forme liaisonnée (94.6%). Dans ces deux dernières tâches, on aurait pu s'attendre à des résultats plus marqués. Nous pensons pouvoir imputer ceux de la tâche de jugement à la tâche elle-même : il s'agissait d'une épreuve longue et dans laquelle les enfants étaient relativement passifs. Ainsi, ils se lassaient assez rapidement et pouvaient fournir des réponses peu réfléchies. Quant aux résultats en lecture à haute voix, ils peuvent, selon nous, être interprétés de deux façons. D'un côté, le taux de réalisation est très fort, ce qui suggère que les LO ne posent que peu de diffi-

cultés aux enfants. D'un autre côté, la réalisation des LO n'est pas systématique – contrairement à ce qu'elle l'est chez l'adulte lors de situations d'entretiens (De Jong, 1994). Ce résultat incite à penser que du fait de sa situation à la frontière de deux mots, même si la liaison est systématique à l'oral, elle nécessite un traitement particulier en lecture chez des enfants.

L'usage des liaisons facultatives varie davantage, et ce dans chacune des trois tâches. Cette tendance générale de distinction entre LO et LF est compatible avec l'hypothèse de notre travail, à savoir que la fréquence à laquelle les liaisons sont entendues réalisées influence la réalisation des liaisons et la représentation que les enfants en ont. Ce résultat était connu en dénomination et en jugement, nous venons de le montrer en tâche de lecture.

Ce travail permet en outre d'affiner cette tendance puisqu'en faisant varier les contextes de LF sur un continuum de fréquence de réalisation entre 0% (après *depuis*) et 99.4% (après *très*), les résultats font clairement apparaître des comportements différents selon les contextes lexicaux à la fois dans la tâche de jugements normatifs et dans la tâche de lecture à haute voix. Par ailleurs, on relève des corrélations positives et significatives entre les taux de réalisation chez les adultes et, d'une part, les taux de jugements en faveur de la forme liaisonnée et, d'autre part, les taux de réalisation dans une tâche de lecture. En d'autres termes, l'influence de l'input oral se retrouve dans les connaissances normatives que les enfants ont de la langue et également dans leur manière de lire à haute voix.

Contrairement à une production orale spontanée, en lecture, le contexte de liaison est repérable et le type de consonne de liaison également (mot1 qui se termine par « n », « s », « t », etc.). On pourrait alors penser que cette donnée facilite et augmente le taux de réalisation des liaisons. Or, dans la tâche de lecture que nous avons proposée, les enfants réalisent globalement moins la liaison que les adultes dans une situation d'entretien. Il semblerait donc que l'influence de l'input oral est plus forte que l'influence de la visibilité de la consonne de liaison. Ce résultat pourra être vérifié par des données dont nous disposons chez des locuteurs adultes (lecteurs experts), dans une tâche de lecture à haute voix similaire puisqu'il est possible que des difficultés de lecture persistent

chez les enfants de notre échantillon. Si les résultats se confirment chez des lecteurs experts, ils fourniront des perspectives intéressantes dans le cadre d'études sur les influences qui affectent la lecture, influences qui peuvent être davantage liées à l'usage de la langue orale qu'à des réalités graphiques.

## 5. Références bibliographiques

- Baclesse, M. (2010). *Influence de la fréquence des liaisons à l'oral sur la lecture chez des enfants dyslexiques*, Mémoire d'orthophonie. Université de Tours.
- Basset, B. (2000). *La liaison à 3, 7 et 11 ans: description et acquisition*. Mémoire de maîtrise, Université Stendhal, Grenoble.
- Bates, E. & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, 12, (5/6), 507-584.
- Blanc, M. & Biggs, P. (1971). L'enquête socio-linguistique sur le français parlé à Orléans. *Le français dans le monde*, 85, 16-25.
- Booij, G. & De Jong, D. (1987). The domain of liaison: theories and data. *Linguistics*, 25, (5), 1005-1025.
- Buben, V. (1935). *Influence de l'orthographe sur la prononciation du français*. Paris : Droz.
- Chabanal, D. & Liégeois, L. (ce volume). Production de liaisons dans l'input parental. In C. Soum Favaro, A. Coquillon, & J.-P. Chevrot (Eds.), *La liaison : approches contemporaines*.
- Chevrot, J.-P., Chabanal, D. & Dugua, C. (2007). Pour un modèle de l'acquisition des liaisons basé sur l'usage: trois études de cas. *Journal of French Language Studies*, 17, 103-128.
- De Jong, D. (1994). La sociophonologie de la liaison orléanaise. In C. Lyche (Ed.), *French Generative Phonology: Retrospective and Perspectives* (pp. 95-129). Salford: ESRI.

- Dugua, C. (2006). *Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage*. Thèse de doctorat, Université Stendhal, Grenoble.
- Dugua, C. (2008). "un ours" / "des ours" ou le rôle de la fréquence sur l'acquisition de la liaison en français. In M. Loiseau, M. Abouzaïd, L. Buson, C. Cavalla, A. Djaroun, C. Dugua, A. Ghimenton, V. Goossens, T. Lebarbé, A. Nardy, F. Rinck & C. Surcouf (Eds.), *Au tour des langues et du langage - Perspective pluridisciplinaire* (pp. 201-208). Grenoble : Presses Universitaires de Grenoble.
- Dugua, C. & Spinelli, E. (2009). La liaison: effets de la fréquence et du rapport à l'écrit sur son acquisition et son usage. In H. Megherbi & M.-C. Fourment (Eds.), *Langages* (pp. 15-40). Paris : L'Harmattan.
- Dugua, C., Spinelli, E., Chevrot, J.-P. & Fayol, M. (2009). Usage-based account of the acquisition of liaison: evidence from sensitivity to plural/singular orientation of nouns. *Journal of Experimental Child Psychology*, 102, 342-350.
- Durand, J. & Lyche, C. (2008). French liaison in the light of corpus data. *Journal of French Language Studies*, 18, (1), 33-66.
- Encrevé, P. (1988). *La liaison avec et sans enchaînement, phonologie tridimensionnelle et usage du français*. Paris : Edition du Seuil.
- Ferrand, L. (2001). *Cognition et lecture. Processus de base de la reconnaissance des mots écrits chez l'adulte*. Bruxelles : De Boeck Université.
- Gaskell, M. G., Spinelli, E. & Meunier, F. (2002). Perception of Resyllabification in French. *Memory and Cognition*, 30, (5), 798-810.
- Grainger, J. & Van Heuven, W. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental Lexicon: Some Word to Talk about Words* (pp. 1-24). Hauppauge, NY: Nova Science Publisher.
- Gunnarsson, C. & Soum Favaro, C. (2008). Liaison et production écrite: l'écrit comme moyen d'étudier un phénomène oral. In D. Alamargot, J. Bouchand, E. Lambert, V. Millogo & C. Beaudet (Eds.), *Proceedings of the International Conference « de la France au Québec : l'écriture dans tous ses états »*. Poitiers.

- Kemmer, S. & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language use* (pp.VII-XXVIII). Stanford Californie: CSLI Publications.
- Méradji, O. & Grégoire, G. (2001). *Les liaisons phonétiques au sein d'une famille*. Mémoire de maitrise, Université Stendhal, Grenoble.
- Nardy, A. (2008). *Acquisition des variables sociolinguistiques entre 2 et 6 ans : facteurs sociologiques et influences des interactions au sein du réseau social*. Thèse de doctorat, Université Stendhal, Grenoble.
- New, B., Pallier, C., Ferrand, L. & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : Lexique. *L'année psychologique*, 101, 447-462.
- Soum-Favaro, C. (2010). De la liaison à l'oral et à l'écrit : approche comparative à l'école primaire. *Colloque Phonlex*. Toulouse, 8-10 septembre 2010.
- Soum Favaro, C., Gunnarsson, C. & Largy, P. (ce volume). La liaison à l'interface entre l'oral et l'écrit. In C. Soum Favaro, A. Coquillon, & J.-P. Chevrot (Eds.), *La liaison : approches contemporaines*.
- Spinelli, E., Cutler, A. & Mcqueen, J. M. (2002). Resolution of liaison for lexical access in French. *Revue Française de Linguistique Appliquée*, VII-1, 83-96.
- Spinelli, E., Mcqueen, J. M. & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language*, 48, 233-254.
- Spinelli, E. & Meunier, F. (2005). Le traitement cognitif de la liaison dans la reconnaissance de la parole enchaînée. *Langages*, 158, 79-88.
- Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, Massachussets: Harvard University Press.
- Yersin-Besson, C. & Grosjean, F. (1996). L'effet de l'enchaînement sur la reconnaissance des mots dans la parole continue. *L'année psychologique*, 96, 9-30.
- Ziegler, J.-C., Ferrand, L. & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory and Cognition*, 32, (5), 732-741.

- Ziegler, J.-C., Muneaux, M. & Grainger, J. (2003). Neighborhood effects in auditory word recognition: phonological competition and orthographic facilitation. *Journal of Memory and Language*, 48, 779-793.



